

WHAT IS CLAIMED IS

5/17  
AZ

1. A method for synthesizing speech with an apparatus comprising a sound source for generating a frequency signal, a vocal tract filter for filtering said frequency signal to generate a speech waveform signal, said filter having characteristics corresponding to a linear predictive coefficient calculated from respective phonemes in a phoneme series, comprising the steps of:

dividing said phonemes into a plurality of frames having a predetermined time length,

summing squares of speech samples in one of said plurality of frames for each frame as a frame power value,

standardizing frame power values at head and tail frames in one phoneme to predetermined values, respectively, to obtain a frame power value of an n-th frame,

summing squares of signal levels of a frame in said frequency signal to obtain a frame power correction value, and

providing a speech envelope signal by means of a function having variables of said standardized frame power values and said frame power correction value, and adjusting an amplitude level of said speech waveform signal as a function of the speech envelope signal.

2. A method according to claim 1, further comprising:

providing power frequency characteristics based on said linear predictive coefficient corresponding to said n-th frame,

calculating an average value of power values sampled from said power frequency characteristics at a predetermined

frequency interval as a mean frame power value,

calculating a speech waveform signal by means of a function having variables of said standardized frame power value, said frame power correction value and said mean frame power value, and

adjusting an amplitude of said speech waveform signal as a function of said speech envelope signal.

3. A method according to claim 2, wherein said function is expressed;

$$V_m = \sqrt{P_n / (G_s G_f)}$$

wherein  $P_n$  is said standardized frame power value,  $G_s$  is said frame power correction value, and  $G_f$  is said mean frame power value.

4. A method according to claim 1, wherein said frequency signal includes an impulse signal carrying a voiced sound and a noise signal carrying an unvoiced sound.